# 产品需求文档（PRD）：AI 图像生成与编辑平台

## 一、产品概述

本产品是一款基于人工智能的图像生成与编辑平台，为用户提供从无到有的文生图、基于图片的混合生成、以及对图像进行局部再创作的全流程工具。核心价值在于用户可以直观地在画布上组合多张图片，输入文字描述，设定风格和比例，利用 AI 模型生成全新的图像，并通过编辑工具进行局部修改和扩展。目标用户包括设计师、创意工作者、运营人员及广大 AI 艺术爱好者。

## 二、项目目标

1. **降低创作门槛**：让没有绘画基础的用户也能通过简单的拖拽、输入文字和选择参考图来生成高质量的图片。
2. **提高生成可控性**：通过多图混合、文本描述、风格参考等参数，使生成结果可控且符合用户预期。
3. **提供完善的编辑能力**：支持局部重绘、扩图、消除笔、细节修复等功能，便于对生成内容进行迭代和优化。
4. **易用且一致的用户体验**：界面清晰直观，操作逻辑统一，支持多端访问和云端同步。

## 三、功能范围

### 3.1 图像上传

* 支持上传 PNG、JPG、WEBP 等常见格式。
* 上传后自动生成新图层，并在画布中显示。
* 上传的图片可拖拽调整位置、缩放大小，支持图层顺序调整。
* 上传入口位于左侧导航栏，图标为带“+”的小图片。

### 3.2 文生图（文本生成图像）

* 通过文字描述创建新的图像，适用于从空白创作或提供文字故事场景。
* 用户需输入描述词（必填，最多 800 字），用以告诉模型想要的内容，例如主体、环境、气氛等。
* 提供多个生成模型供选择，模型说明展示其擅长风格或效果。
* 可上传一张风格参考图，模型会在生成时参考其色调和画风；如不选择则按默认风格生成。
* 支持自定义输出尺寸及常见宽高比例的快捷选择（1:1、4:3、3:2、16:9、21:9 等）。
* 点击“立即生成”后启动生成，生成的图像会保存到生成记录并在画布中作为新图层显示。

### 3.3 图生图（基于图片生成图片）

* 允许用户选取画布上的多张图片作为参考，生成一张新的融合图像。
* 提供“混合图层预览”区域显示参与混合的图片缩略图，用户可移除或添加图层。
* 文本描述词为可选，用于补充或引导生成意图。
* 参考强度可调节，用滑块控制生成结果对参考图的依赖程度，数值越高生成越接近原图。
* 高级设置中可上传风格图，并选择“混合参考”或“分图层参考”模式：
* 混合参考：系统整体参照所有图层。
* 分图层参考：可分别设置不同图层的参考权重或特征维度。
* 细分参考维度选项包括：轮廓边缘、景深、人物姿势、图片信息（色彩、光影）、主体等。用户可根据需要勾选。
* 点击“立即生成”触发生成，结果保存至生成记录和画布。

### 3.4 编辑与再创作工具

系统提供的编辑工具位于顶部二级工具栏，用于对画布上选中的区域或图层进行二次创作：

* **局部重绘**：选择区域后重新生成该部分内容，可输入提示词或直接使用智能重绘，适用于修改画面局部或替换物体。
* **扩图**：在原有图像基础上向外延伸场景，用户选择扩展方向和区域，输入提示词，模型会保持风格一致性进行扩充。
* **消除笔**：用于删除图片中不需要的元素，划定区域后系统分析周围背景并填充删除区域。
* **细节修复**：提升局部区域的细节和质感，修复模糊或噪点。
* **HD 超清**：对整张图片进行超分辨率处理，提升清晰度和像素密度。
* **智能超清**：基于 AI 的细节优化，针对不同区域（如人物、人脸、背景等）采取差异化增强策略。
* **抠图**：自动识别并分离前景主体，生成透明背景图片或新的图层，方便后期合成。

### 3.5 图层与生成记录管理

* **图层管理**：右侧“图层”面板列出所有图层，支持点击选中、拖拽调整顺序。未来版本可增加锁定、隐藏、复制等操作。
* **生成记录**：在“生图记录”标签下保存所有生成历史，用户可预览、重新插入画布或删除不需要的记录。
* **图层与记录的可视化**：点击缩略图高亮对应图层，便于定位；记录列表支持分页或滚动加载。

### 3.6 画布与项目管理

* 实时画布开关：支持实时渲染和预览模式切换，实时模式下所有操作即时呈现，预览模式用于低性能设备或大规模场景。
* 画布尺寸调节：顶栏中显示当前画布大小，点击可选择预设尺寸或自定义输入；调整后原有图层按比例缩放。
* 选择/手形工具：选择工具用于选中和调整单个图层；手形工具用于拖动画布整体视图。
* 文本工具：在画布中插入文本对象，可调节字体、字号、颜色等。
* 画笔工具：提供自由绘制或标记功能，为局部重绘划定范围或进行草图标注。
* 撤销/重做：支持逐步撤回和恢复操作。
* 保存与导出：项目自动保存至云端，用户可通过“导出”按钮导出当前画布为 PNG/JPG，并选择分辨率或透明背景。

## 四、用户故事

1. **初学者生成场景图**：小白用户希望根据一段文字描述生成一个卡通风格的公园场景。他打开文生图面板，输入描述词“明亮的公园，秋千、滑梯和正在玩耍的小朋友”，选取1:1比例并上传参考风格图，点击生成后得到一张符合预期的图像。
2. **设计师混合素材**：平面设计师上传多张参考图（人物、背景、装饰），通过图生图功能设置参考强度，并在高级设置中只勾选“人物姿势”和“主体”，生成一张人物与背景融合且姿态自然的新图。随后使用局部重绘修饰面部细节。
3. **运营人员制作海报**：运营人员在画布上上传产品图片，使用抠图分离主体，再用扩图功能补全背景；随后添加文本层写入促销信息，通过 HD 超清提升整体画质，最后导出高分辨率海报图用于线上投放。

## 五、非功能需求

* **性能要求**：普通生成任务平均响应时间在 10 秒以内；图生图与超清处理不超过 30 秒。
* **可靠性**：自动保存机制确保编辑过程不会因为意外中断而丢失数据；生成记录至少保留最近 30 条历史。
* **易用性**：界面层级清晰，文字描述、按钮和图标均应易懂；提供新手引导和帮助文档。
* **兼容性**：支持主流现代浏览器，针对不同分辨率自动适配；后续计划推出移动端应用。
* **隐私与安全**：上传的图片和生成内容仅在用户个人账户中可见，支持加密存储和传输，不会对外泄漏。

## 六、技术实施建议

* **AI 模型集成**：支持多模型切换，文生图和图生图分别调用不同模型 API，根据用户选择按需加载。
* **图层与画布引擎**：使用 HTML5 Canvas 或 WebGL 实现，支持多图层、自由拖拽和缩放的高性能渲染。
* **服务器架构**：生成请求异步排队处理，利用微服务或函数式架构优化扩展性。
* **前端框架**：建议使用现代框架（如 React/Vue）管理界面状态；结合状态管理（Redux/Pinia）组织数据。
* **云端存储**：使用对象存储（如 AWS S3、阿里云 OSS）保存图片文件，数据库存储用户配置、生成记录等元数据。

## 七、风险与挑战

* **生成结果的不可控性**：即使设置了参考图和描述，AI 生成结果仍具有随机性，需要提供足够的迭代手段（如局部重绘、参考强度调节）增加可控性。
* **版权与审核**：需遵守版权政策，限制上传受版权保护的图像和生成敏感内容；后台加入审核机制。
* **性能瓶颈**：图生图和超清处理消耗算力较大，要做好队列控制和并发管理；必要时采用 CDN 加速图像下载。

## 八、后续迭代方向

* 引入更多专业模型与风格包，丰富生成效果。
* 提升图层管理功能，如分组、锁定、调整透明度等。
* 扩展到移动端，支持触控操作与离线编辑。
* 与第三方图库、社交平台打通，实现一键导入素材和分享成果。

## 九、个人建议

作为产品运营，对这一平台的定位和功能组合持肯定态度。文生图与图生图的有机结合既满足了用户从零创作的需求，也能够为已有素材注入新的灵感。建议在后续迭代中重点优化用户反馈机制，通过 A/B 测试收集不同参数组合的满意度，用数据指导模型调优。同时，可在新手引导中加入案例教程，让用户快速掌握参考强度、风格参考的使用技巧，从而提升平台粘性。